

Polynomial-sized Semidefinite Representations of Derivative Relaxations of Spectrahedral Cones

James Saunderson

Pablo A. Parrilo*

August 8, 2012

Abstract

We give explicit polynomial-sized (in n and k) semidefinite representations of the hyperbolicity cones associated with the elementary symmetric polynomials of degree k in n variables. These convex cones form a family of non-polyhedral outer approximations of the non-negative orthant that preserve low-dimensional facial structure while successively discarding high-dimensional facial structure.

More generally we construct explicit semidefinite representations (polynomial sized in k, m , and n) of the hyperbolicity cones associated with k th directional derivatives of polynomials of the form $p(x) = \det(\sum_{i=1}^n A_i x_i)$ where the A_i are $m \times m$ symmetric matrices. These convex cones form an analogous family of outer approximations to any spectrahedral cone.

Our representations allow us to use semidefinite programming to efficiently solve the linear cone programs associated with these convex cones as well as their (less well understood) dual cones.

1 Introduction

Expressing convex optimization problems in conic form, as the minimization of a linear functional over an affine slice of a convex cone, has been an important method in the development of modern convex optimization theory. This abstraction is useful (at least from a theoretical viewpoint) because all that is difficult and interesting about the problem is packaged into the cone. The conic viewpoint provides a natural way to organize classes of convex optimization problems into hierarchies based on whether the cones associated with one class can be expressed in terms the cones associated with another class. For example, semidefinite programming generalizes linear programming because the non-negative orthant is the restriction to the diagonal of the positive semidefinite cone.

When faced with a convex cone the geometry of which is not well understood, we stand to gain theoretical insight as well as off-the-shelf optimization algorithms by representing it in terms of a cone with known geometric and algebraic structure such as the positive semidefinite cone. Terminology is attached to this idea, with a cone being *spectrahedral* if it is a linear section of the positive semidefinite cone, and *semidefinitely representable* if it is a linear projection of a spectrahedral cone. The efficiency of a semidefinite representation is also clearly important. If we can write a cone as the projection of a slice of the cone of $m \times m$ positive semidefinite matrices,

*The authors are with the Laboratory for Information and Decision Systems, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge MA 02139, USA. Email: {james.s, parrilo}@mit.edu. This research was funded by the Air Force Office of Scientific Research under grants FA9550-11-1-0305 and FA9550-12-1-0287.

we say it has a semidefinite representation of *size* m . Many convex cones have been shown to be semidefinitely representable using a variety of techniques (see [4] and [13] for contrasting methods and examples).

The classes of semidefinitely representable cones and spectrahedral cones are distinct, although semidefinitely representable cones are perhaps more natural from the point of view of optimization. Unlike spectrahedral cones, the class of semidefinitely representable cones is closed under duality. Furthermore, a semidefinite representation of a cone suffices to express the associated cone program as a semidefinite program.

The *hyperbolicity cones* form a large family of convex cones (constructed from certain multivariate polynomials) that includes the positive semidefinite cone, as well as all homogeneous cones. While it has been shown (by Lewis et al. [11] based on work of Helton and Vinnikov [8]) that all three-dimensional hyperbolicity cones are spectrahedral, little is known about semidefinite representations of higher dimensional hyperbolicity cones. Furthermore while hyperbolicity cones have very simple descriptions, their dual cones are not well understood.

In this paper we give explicit, polynomial-sized semidefinite representations of the hyperbolicity cones associated with the elementary symmetric polynomials. These cones form a family of non-polyhedral outer approximations to the non-negative orthant known as *derivative relaxations*. Our semidefinite representations naturally generalize to analogous families of derivative relaxations of spectrahedral cones.

1.1 Hyperbolic polynomials and hyperbolicity cones

A homogeneous polynomial p of degree m in n variables is *hyperbolic* with respect to $e \in \mathbb{R}^n$ if $p(e) \neq 0$ and if for all $x \in \mathbb{R}^n$ the univariate polynomial $t \mapsto p(x - te)$ has only real roots. These m real roots, denoted $\lambda_e(x) \in \mathbb{R}^m$, are called the *characteristic roots* of x (with respect to e). Gårding's foundational work on hyperbolic polynomials [6] establishes that if p is hyperbolic with respect to e then the connected component of $\{x \in \mathbb{R}^n : p(x) \neq 0\}$ containing e is an open convex cone. This cone is called the *hyperbolicity cone* corresponding to (p, e) . We denote it by $\Lambda_{++}(p, e)$, and its closure by $\Lambda_+(p, e)$.

Since p is hyperbolic with respect to e if and only if $-p$ is hyperbolic with respect to e , we assume, throughout, that $p(e) > 0$.

Basic examples:

- The polynomial $p(x_1, x_2, \dots, x_n) = x_1 x_2 \cdots x_n$ is hyperbolic with respect to $e = \mathbf{1}_n := (1, 1, \dots, 1)$. The vector of characteristic roots is $\lambda_e(x) = x$. The associated closed hyperbolicity cone is the non-negative orthant.
- Let A_1, A_2, \dots, A_n be symmetric $m \times m$ matrices and $e \in \mathbb{R}^n$ such that $\sum_{i=1}^n e_i A_i$ is positive definite. The closed hyperbolicity cone associated with the polynomial $p(x_1, \dots, x_n) = \det(\sum_{i=1}^n x_i A_i)$ and the direction e is the spectrahedral cone $\Lambda_+(p, e) = \{x \in \mathbb{R}^n : \sum_{i=1}^n x_i A_i \succeq 0\}$.

Roots vs coefficients: Throughout this paper we often pass back and forth between root-based and coefficient-based views of hyperbolic polynomials and their hyperbolicity cones. Indeed if p has degree m and is hyperbolic with respect to $e \in \mathbb{R}^n$ then the following two equivalent conditions characterize $x \in \Lambda_+(p, e)$:

- the roots of $t \mapsto p(x - te)$ are all non-negative, i.e.

$$\Lambda_+(p, e) = \{x \in \mathbb{R}^n : \lambda_e(x) \in \mathbb{R}_+^m\} \quad (1)$$

- the coefficients of $t \mapsto p(x + te) = p(e)t^m + \sum_{i=1}^m a_i(x)t^{m-i}$ are all non-negative, i.e.

$$\Lambda_+(p, e) = \{x \in \mathbb{R}^n : a_1(x) \geq 0, \ a_2(x) \geq 0, \ \dots, \ a_m(x) \geq 0\}. \quad (2)$$

This equivalence is the specialization of Descartes' classical 'rule of signs' to the case of polynomials with real roots [15, Theorem 20]. Note that the coefficients $a_i(x)$ appearing in (2) are themselves homogeneous polynomials of degree i in x .

1.2 Derivative relaxations

If p is hyperbolic with respect to e then (essentially by Rolle's theorem) the directional derivative of p in the direction e , viz.

$$p_e^{(1)}(x) := \left. \frac{d}{dt} p(x + te) \right|_{t=0}$$

is also hyperbolic with respect to e , a construction that goes back to Gårding [6]. If p has degree m , by repeatedly differentiating in the direction e we construct a sequence of polynomials $p, p_e^{(1)}, p_e^{(2)}, \dots, p_e^{(m-1)}$ each hyperbolic with respect to e . In terms of coefficients, if $p(x + te) = p(e)t^m + \sum_{i=1}^m a_i(x)t^{m-i}$ then the k th derivative essentially removes high degree terms, i.e.

$$p_e^{(k)}(x + te) = c_0 a_{m-k}(x) + c_1 a_{m-k-1}(x)t + \dots + c_{m-k} p(e)t^{m-k}$$

where $c_i = (k+i)!/i! > 0$. Hence the corresponding hyperbolicity cone is

$$\Lambda_+(p_e^{(k)}, e) = \{x \in \mathbb{R}^m : a_1(x) \geq 0, \ a_2(x) \geq 0, \ \dots, \ a_{m-k}(x) \geq 0\}$$

and can be obtained from (2), the coefficient-based description of $\Lambda_+(p, e)$, by removing k of the inequality constraints. As a result, the hyperbolicity cones $\Lambda_+(p_e^{(k)}, e)$, which we abbreviate by $\Lambda_+^{(k)}(p, e)$, provide a sequence of outer approximations to the original hyperbolicity cone that satisfy

$$\Lambda_+(p, e) \subset \Lambda_+^{(1)}(p, e) \subset \dots \subset \Lambda_+^{(m-1)}(p, e).$$

The last of these, $\Lambda_+^{(m-1)}(p, e)$, is simply the closed half-space defined by e . The work of Renegar [15] highlights the many nice properties of this sequence of approximations.

Note that we abuse terminology by referring to the cones $\Lambda_+^{(k)}(p, e)$ as *derivative relaxations* of the hyperbolicity cone $\Lambda_+(p, e)$. The abuse is that $\Lambda_+^{(k)}(p, e)$ does not depend only on the *geometric* object $\Lambda_+(p, e)$ but on its particular *algebraic* description via p and e .

1.3 Derivative relaxations of the non-negative orthant

In the case of $p(x) = x_1 x_2 \dots x_n$ and $e = \mathbf{1}_n$, we have that

$$p(x + te) = \sigma_n(x) + \sigma_{n-1}(x)t + \dots + \sigma_1(x)t^{n-1} + t^n$$

where $\sigma_k(x) = \sum_{1 \leq i_1 < \dots < i_k \leq n} x_{i_1} \cdots x_{i_k}$ denotes the k th elementary symmetric polynomial in n variables. As such the k th derivative relaxation of the non-negative orthant, for which we reserve the special notation $\mathbb{R}_+^{n,(k)}$, has the following two equivalent descriptions:

$$\mathbb{R}_+^{n,(k)} = \{x \in \mathbb{R}^n : \text{all the roots of } t \mapsto \frac{d^k}{dt^k} \sigma_n^n(x - te) \text{ are non-negative}\} \quad (3)$$

$$\mathbb{R}_+^{n,(k)} = \{x \in \mathbb{R}^n : \sigma_1(x) \geq 0, \sigma_2(x) \geq 0, \dots, \sigma_{n-k}(x) \geq 0\}. \quad (4)$$

Consistent with these descriptions we define $\mathbb{R}_+^{n,(n)} := \mathbb{R}^n$.

We emphasise the cones $\mathbb{R}_+^{n,(k)}$ because the derivative relaxations of any hyperbolicity cone can be expressed in terms of the derivative relaxations of the orthant via the following (an immediate consequence of [15, Proposition 18]).

Proposition 1. *If p has degree m and is hyperbolic with respect to $e \in \mathbb{R}^n$ then*

$$\Lambda_+^{(k)}(p, e) = \left\{x \in \mathbb{R}^n : \lambda_e(x) \in \mathbb{R}_+^{m,(k)}\right\}. \quad (5)$$

Thus, having a semidefinite representation of $\mathbb{R}_+^{n,(k)}$ together with a semidefinite characterization (in some appropriate sense) of $\lambda_e(x)$, allows us to construct a semidefinite representation for $\Lambda_+^{(k)}(p, e)$.

1.4 Results

In this paper we construct two different explicit semidefinite representations of the derivative relaxations of the non-negative orthant. Our *root-based representation* (see Section 3.1) is based on the characterization of $\mathbb{R}_+^{n,(k)}$ in (3). Our *coefficient-based representation* (see Section 3.2) is based on the characterization of $\mathbb{R}_+^{n,(k)}$ in (4). Our constructions establish the following result.

Theorem 1. *The derivative relaxations of the non-negative orthant have semidefinite representations of polynomial size, i.e. for each $n \in \mathbb{N}$ and $1 \leq k \leq n$, the cone $\mathbb{R}_+^{n,(k)}$ is a projection of a slice of the cone of $m(n, k) \times m(n, k)$ positive semidefinite matrices where $m(n, k) \in O(n^2 \min\{k, n - k\})$.*

While both of our representations of $\mathbb{R}_+^{n,(k)}$ are polynomial sized in n and k , the root-based representation is smaller when k is small while the coefficient-based representation is smaller when $n - k$ is small. (Hence the dependence on $\min\{k, n - k\}$ in the statement of Theorem 1.)

Using Proposition 1 it is fairly straightforward to generalize these results to construct semidefinite representations (of size polynomial in k, m , and n) of the cones $\Lambda_+^{(k)}(p, e)$ where $p(x) = \det(\sum_{i=1}^n A_i x_i)$, the A_1, \dots, A_n are symmetric $m \times m$ matrices, and e is a direction of hyperbolicity for p . We discuss this in more detail in Section 4, and present an example illustrating a construction of this type.

1.5 Related work

Previous work has also focused on semidefinite representations of the derivative relaxations of the orthant. Zinchenko [20] used a decomposition approach to give semidefinite representations of $\mathbb{R}_+^{n,(1)}$ and its dual cone. Sanyal [17] subsequently gave spectrahedral representations of $\mathbb{R}_+^{n,(1)}$ and $\mathbb{R}_+^{n,(n-2)}$ and conjectured that all of the derivative relaxations of the orthant admit spectrahedral representations. Recently Brändén [2] settled this conjecture, giving representations of each of

the cones $\mathbb{R}^{n,(k)}$ for $k = 1, 2, \dots, n-1$ as the intersection of a subspace with the cone of $m \times m$ positive semidefinite matrices for $m \sim n^{n-k}$. While this is of considerable theoretical interest, these representations (unlike ours) are not practical for optimization due to their prohibitive size.

The spectrahedral representation of $\mathbb{R}_+^{n,(1)}$ given by Sanyal was already implicitly given in the work of Choe et al. [3] that studies the relationships between matroids and hyperbolic polynomials. Choe et al. observe that if \mathcal{M} is a *regular* matroid represented by the rows of a totally unimodular matrix V then $\det(V^T \text{diag}(x)V)$ is the basis generating polynomial of \mathcal{M} . In particular, the uniform matroid U_n^{n-1} is regular and has $\sigma_{n-1}(x)$ as its basis generating polynomial, yielding a symmetric determinantal representation of $\sigma_{n-1}(x)$ and hence a spectrahedral representation of $\mathbb{R}_+^{n,(n-1)}$.

From a computational perspective, Güler [7] showed that if p has degree m and is hyperbolic with respect to e then $\log p$ is a self-concordant barrier function (with barrier parameter m) for the hyperbolicity cone $\Lambda_+(p, e)$. As such, as long as p and its gradient and Hessian can be computed efficiently, one can use interior point methods to minimize a linear functional over an affine slice of $\Lambda_+(p, e)$ efficiently. Renegar [15, Section 9] gave an efficient interpolation-based method for computing $p_e^{(k)}$ (and its gradient and Hessian) whenever p (and its gradient and Hessian) can be evaluated efficiently. Güler and Renegar's observations together yield efficient computational methods to optimize a linear functional over an affine slice of a derivative relaxation of a spectrahedral cone. Our results complement these, giving an efficient method to solve optimization problems of this type using existing numerical procedures for semidefinite programming.

1.6 A semidefinite representation of $\mathbb{R}_+^{5,(2)}$

As a concrete example of our representations, we give the coefficient-based representation of $\mathbb{R}_+^{5,(2)}$: the first case not covered by the previous work on this topic (excluding Brändén's).

Example 1. For each n let $Q_n = [\mathbf{1}_n/\sqrt{n} \ V_n]$ denote an $n \times n$ orthogonal matrix with first column $(1, 1, \dots, 1)/\sqrt{n}$. Then

$$\begin{aligned} x \in \mathbb{R}_+^{5,(5-3)} &\iff \exists y^{(1)} \in \mathbb{R}_+^4, \ z_2^{(1)}, z_3^{(1)} \in \mathbb{R}, \ M^{(1)}, Z_2^{(1)}, Z_3^{(1)} \in \mathcal{S}_+^4 \\ &\quad \exists y^{(2)} \in \mathbb{R}_+^3, \ z_2^{(2)} \in \mathbb{R}, \ M^{(2)}, Z_2^{(2)} \in \mathcal{S}_+^3 \end{aligned}$$

such that

$$\begin{aligned} Q_5^T \text{diag}(x) Q_5 &\succeq \begin{bmatrix} 0 & 0 \\ 0 & M^{(1)} \end{bmatrix} & Q_4^T \text{diag}(y^{(1)}) Q_4 &\succeq \begin{bmatrix} 0 & 0 \\ 0 & M^{(2)} \end{bmatrix} \\ \text{tr}(M^{(1)}) &= \sum_{j=1}^4 y_j^{(1)} & \text{tr}(M^{(2)}) &= \sum_{j=1}^3 y_j^{(2)} \\ M^{(1)} &\succeq y_1^{(1)} I & M^{(2)} &\succeq y_1^{(2)} I \\ \text{for } i = 2, 3: & & \text{for } i = 2: & \\ M^{(1)} &\succeq z_i^{(1)} I - Z_i^{(1)} & M^{(2)} &\succeq z_i^{(2)} I - Z_i^{(2)} \\ iz_i^{(1)} - \text{tr}(Z_i^{(1)}) &\geq \sum_{j=1}^i y_j^{(1)} & iz_i^{(2)} - \text{tr}(Z_i^{(2)}) &\geq \sum_{j=1}^i y_j^{(2)} \end{aligned}$$

and

$$y_1^{(2)} + y_2^{(2)} + y_3^{(2)} \geq 0.$$

The representation as presented here looks quite complicated only because, unlike elsewhere in the paper, we have chosen to write it out rather explicitly, ignoring simplifications that can be made (see Section 5) and abstractions (see Section 2) that allow for more compact expression of the same representation.

1.7 Notation

Here we define notation not explicitly defined elsewhere in the paper.

Sets: Let \mathcal{S}^n and \mathcal{S}_+^n denote the sets of $n \times n$ symmetric and symmetric positive semidefinite matrices respectively. We write $X \succeq Y$ if $X - Y \in \mathcal{S}_+^n$ for some n . We denote the cone of vectors in non-decreasing order by $\mathbb{R}_\uparrow^n = \{x \in \mathbb{R}^n : x_1 \leq x_2 \leq \dots \leq x_n\}$, and write $x_{(i)}$ for the i th smallest entry of $x \in \mathbb{R}^n$. If C is a convex cone, we denote by C^* the dual cone, i.e. the set of linear functionals that are non-negative on C . We represent linear functionals on \mathbb{R}^n using the standard Euclidean inner product, and linear functionals on \mathcal{S}^n using the trace inner product $\langle X, Y \rangle = \text{tr}(XY)$. Finally, the closure of a set S is denoted $\text{cl}(S)$.

Maps: If $X \in \mathcal{S}^n$ let $\lambda(X) \in \mathbb{R}_\uparrow^n$ denote the sorted eigenvalues of X . If $x \in \mathbb{R}^n$ then $\text{diag}(x) \in \mathcal{S}^n$ is the symmetric matrix with diagonal entries given by x . If $X \in \mathcal{S}^n$ then $\text{diag}(X) \in \mathbb{R}^n$ is the vector of diagonal elements of X . The usage will be clear from the context.

2 Background on spectral constraints and semidefinite representations

2.1 Spectral constraints

Our semidefinite representations are built upon semidefinite representations for convex cones (and their dual cones) of the form

$$\lambda^{-1}[C] := \{X \in \mathcal{S}^n : \lambda(X) \in C\}$$

where $C \subset \mathbb{R}^n$ is a closed semidefinitely representable cone that is invariant under permutations of the coordinates (here we call this a *permutation invariant* convex cone). We also require semidefinite representations of convex cones of the form

$$\begin{aligned} \lambda_{sc}^{-1}[C] &:= \text{cl} \left\{ \begin{bmatrix} t & u^T \\ u & X \end{bmatrix} \in \mathcal{S}^{n+1} : t > 0, \ X - uu^T/t \in \lambda^{-1}[C] \right\} \\ &= \left\{ \begin{bmatrix} t & u^T \\ u & X \end{bmatrix} \in \mathcal{S}^{n+1} : t \geq 0, \ X \in \lambda^{-1}[C], \ tX - uu^T \in \lambda^{-1}[C] \right\} \end{aligned}$$

where $C \supseteq \mathbb{R}_+^n$ is a closed semidefinitely representable permutation invariant cone containing the non-negative orthant. Observe that $X - uu^T/t$ is the *Schur complement* of t in the partitioned matrix $\begin{bmatrix} t & u^T \\ u & X \end{bmatrix}$ and that the cone $\lambda_{sc}^{-1}[C]$ is the *Siegel cone* associated with $\lambda^{-1}[C]$ and the bilinear form $B(u, v) = (uv^T + vu^T)/2$. (For a discussion of Siegel cones and their relationship with homogeneous cones see [1], for example.)

2.2 Majorization

If $x, y \in \mathbb{R}^n$ we write $x \trianglelefteq y$ if x is *majorized* by y , i.e.

$$\sum_{i=1}^n x_{(i)} = \sum_{i=1}^n y_{(i)} \quad \text{and} \quad \sum_{i=1}^k x_{(i)} \geq \sum_{i=1}^k y_{(i)} \quad \text{for } k = 1, 2, \dots, n-1$$

where $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$ denotes the entries of x sorted in ascending order. Geometrically $x \trianglelefteq y$ if and only if x is in the convex hull of the vectors obtained by permuting the entries of y .

It is well known that $\lambda^{-1}[C]$ can be rewritten as

$$\lambda^{-1}[C] = \{X \in \mathcal{S}^n : \exists y \in C \text{ such that } \lambda(X) \trianglelefteq y\}. \quad (6)$$

To see this, on the one hand note that if $\lambda(X) \in C$ we can take $y = \lambda(X)$. On the other hand if $y \in C$ then the convex hull of the permutations of the entries of y is also in C (by permutation invariance and convexity of C) and so if $\lambda(X) \trianglelefteq y$ it follows that $\lambda(X) \in C$.

2.3 Explicit semidefinite representations of $\lambda^{-1}[C]$ and $\lambda_{sc}^{-1}[C]$

A semidefinite representation of sets of the form $\lambda^{-1}[C]$ is given in [13, Section 4.3.1]. We now present a slight variation that is more suited to our purposes.

Proposition 2. *If $C \subset \mathbb{R}^n$ is a permutation invariant convex cone then $X \in \lambda^{-1}[C]$ if and only if*

$$\begin{aligned} &\text{there exists } y \in C \cap \mathbb{R}_+^n, \quad z_2, \dots, z_{n-1} \in \mathbb{R}, \quad Z_2, \dots, Z_{n-1} \succeq 0 \\ &\text{such that } \operatorname{tr}(X) = \sum_{j=1}^n y_j, \quad X \succeq y_1 I, \quad \text{and} \end{aligned} \quad (7)$$

$$\text{for } i = 2, \dots, n-1, \quad X \succeq z_i I - Z_i \quad \text{and} \quad iz_i - \operatorname{tr}(Z_i) \geq \sum_{j=1}^i y_j. \quad (8)$$

Proof. Suppose such y , z_i and Z_i exist (for $i = 2, 3, \dots, n-1$). In light of (6) we need to show that $\lambda(X) \trianglelefteq y$. From (7) we have that $\sum_{j=1}^n \lambda_{(j)}(X) = \operatorname{tr}(X) = \sum_{j=1}^n y_{(j)}$ and $\lambda_{(1)}(X) \geq y_{(1)}$. Furthermore, since $X \succeq z_i I - Z_i$ and $Z_i \succeq 0$, the sum of the smallest i eigenvalues (for $i = 2, 3, \dots, n-1$) of X satisfies

$$\sum_{j=1}^i \lambda_{(j)}(X) \geq \sum_{j=1}^i \lambda_{(j)}(z_i I - Z_i) = iz_i - \sum_{j=1}^i \lambda_{(j)}(Z_i) \geq iz_i - \operatorname{tr}(Z_i) \geq \sum_{j=1}^i y_{(j)}$$

where the last inequality is from (8). Hence $\lambda(X) \trianglelefteq y$ and so $X \in \lambda^{-1}[C]$.

For the other direction if $X \in \lambda^{-1}[C]$ it is straightforward to check that we can satisfy (7) and (8) by choosing $y = \lambda(X)$, $z_i = y_i$ for $i = 2, \dots, n-1$, and

$$Z_i = \sum_{j=1}^{i-1} [\lambda_{(i)}(X) - \lambda_{(j)}(X)] u_j u_j^T$$

where u_j is an eigenvector of X corresponding to $\lambda_{(j)}(X)$ and the u_j are orthonormal. \square

When C is a permutation invariant convex cone containing the non-negative orthant we can give an explicit semidefinite representation of $\lambda_{sc}^{-1}[C]$ using our representation for $\lambda^{-1}[C]$ together with the following observation.

Proposition 3. *If $C \subset \mathbb{R}^n$ is a permutation invariant convex cone such that $C \supset \mathbb{R}_+^n$ then*

$$\lambda_{sc}^{-1}[C] = \left\{ \begin{bmatrix} t & u^T \\ u & X \end{bmatrix} \in \mathcal{S}^{n+1} : \exists M \in \lambda^{-1}[C] \text{ s.t. } \begin{bmatrix} t & u^T \\ u & X \end{bmatrix} \succeq \begin{bmatrix} 0 & 0 \\ 0 & M \end{bmatrix} \right\}. \quad (9)$$

Proof. First we show that $\lambda_{sc}^{-1}[C]$ is contained in the right hand side of (9). Suppose $\begin{bmatrix} t & u^T \\ u & X \end{bmatrix} \in \lambda_{sc}^{-1}[C]$. If $t > 0$ we can take $M = X - uu^T/t \in \lambda^{-1}[C]$. Since $X - M - uu^T/t = 0$ it follows by basic properties of the Schur complement that $\begin{bmatrix} t & u^T \\ u & X - M \end{bmatrix} \succeq 0$. If $t = 0$ and $\begin{bmatrix} t & u^T \\ u & X \end{bmatrix} = \begin{bmatrix} 0 & u^T \\ u & X \end{bmatrix} \in \lambda_{sc}^{-1}[C]$ it follows that $u = 0$ and it suffices to take $M = X \in \lambda^{-1}[C]$.

For the reverse inclusion suppose $t > 0$ and there exists $M \in \lambda^{-1}[C]$ such that $\begin{bmatrix} t & u^T \\ u & X - M \end{bmatrix} \succeq 0$. It follows (again from properties of the Schur complement) that $X - uu^T/t \succeq M$. Since $C \supset \mathbb{R}_+^n$ implies $\lambda^{-1}[C] \supset \mathcal{S}_+^n$ we can conclude that $X - uu^T/t \in \lambda^{-1}[C]$ as required. Finally if $t = 0$ and there exists $M \in \lambda^{-1}[C]$ such that $\begin{bmatrix} t & u^T \\ u & X - M \end{bmatrix} = \begin{bmatrix} 0 & u^T \\ u & X - M \end{bmatrix} \succeq 0$ then $u = 0$ and $X \succeq M$. Hence $tX - uu^T = 0 \in \lambda^{-1}[C]$ and $X \in \lambda^{-1}[C]$ as required. \square

2.4 Duality and spectral constraints

In order to construct semidefinite representations of the dual cones $(\mathbb{R}_+^{n,(k)})^*$ we need the following duality relation for sets of the form $\lambda^{-1}[C]$ where $C \subset \mathbb{R}^n$ is a permutation invariant convex cone. This is a special case of Lewis' result on the convex conjugates of convex spectral functions [10, Theorem 2.3].

Proposition 4. *If $C \subset \mathbb{R}^n$ is a permutation invariant convex cone then*

$$(\lambda^{-1}[C])^* = \lambda^{-1}[C^*] = \{X \in \mathcal{S}^n : \lambda(X) \in C^*\}.$$

The analogous result for $\lambda_{sc}^{-1}[C]$ is as follows.

Proposition 5. *If $C \subset \mathbb{R}^n$ is a permutation invariant convex cone such that $C \supset \mathbb{R}_+^n$ then*

$$(\lambda_{sc}^{-1}[C])^* = \left\{ \begin{bmatrix} s & v^T \\ v & Y \end{bmatrix} \in \mathcal{S}_+^{n+1} : Y \in \lambda^{-1}[C^*] \right\}. \quad (10)$$

Proof. First we show that $(\lambda_{sc}^{-1}[C])^*$ is contained in the right hand side of (10). Since $C \supset \mathbb{R}_+^n$ it follows that $\lambda_{sc}^{-1}[C] \supset \mathcal{S}_+^{n+1}$ and so that $(\lambda_{sc}^{-1}[C])^* \subset (\mathcal{S}_+^{n+1})^* = \mathcal{S}_+^{n+1}$. Now suppose that $\begin{bmatrix} s & v^T \\ v & Y \end{bmatrix} \in (\lambda_{sc}^{-1}[C])^*$. Then for any $X \in \lambda^{-1}[C]$ we have that $\begin{bmatrix} 0 & 0 \\ 0 & X \end{bmatrix} \in \lambda_{sc}^{-1}[C]$ and so that

$$\langle Y, X \rangle = \left\langle \begin{bmatrix} s & v^T \\ v & Y \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & X \end{bmatrix} \right\rangle \geq 0.$$

Hence $Y \in (\lambda^{-1}[C])^* = \lambda^{-1}[C^*]$ by Proposition 4.

For the reverse inclusion assume that $\begin{bmatrix} s & v^T \\ v & Y \end{bmatrix} \succeq 0$ and $Y \in \lambda^{-1}[C^*]$. Let $\begin{bmatrix} t & u^T \\ u & X \end{bmatrix}$ be an arbitrary element of $\lambda_{sc}^{-1}[C]$. By Proposition 3 there is $M \in \lambda^{-1}[C]$ such that $\begin{bmatrix} t & u^T \\ u & X - M \end{bmatrix} \succeq 0$. Then

$$\left\langle \begin{bmatrix} s & v^T \\ v & Y \end{bmatrix}, \begin{bmatrix} t & u^T \\ u & X \end{bmatrix} \right\rangle = \left\langle \begin{bmatrix} s & v^T \\ v & Y \end{bmatrix}, \begin{bmatrix} t & u^T \\ u & X - M \end{bmatrix} \right\rangle + \langle Y, M \rangle \geq 0$$

and so $\begin{bmatrix} s & v^T \\ v & Y \end{bmatrix} \in (\lambda_{sc}^{-1}[C])^*$. \square

3 Semidefinite representations of $\mathbb{R}_+^{n,(k)}$

In this section we present our two semidefinite representations for the cones $\mathbb{R}_+^{n,(k)}$ for $k = 1, 2, \dots, n-1$ corresponding to the root-based and coefficient-based characterizations of $\mathbb{R}_+^{n,(k)}$ in (3) and (4) respectively.

Our two representations are recursive in nature. The root-based representation expresses $\mathbb{R}_+^{n,(k)}$ in terms of $\mathbb{R}_+^{n-1,(k-1)}$, working towards the base case $\mathbb{R}_+^{n-k,(0)}$, i.e. the non-negative orthant. The coefficient-based representation expresses $\mathbb{R}_+^{n,(n-k)}$ in terms of $\mathbb{R}_+^{n-1,(n-k)}$, working towards the base case $\mathbb{R}_+^{n-k+1,(n-k)}$, a half-space.

3.1 Root-based representation

If $x \in \mathbb{R}^n$ define

$$\eta(x) := \text{roots of } t \mapsto \frac{d}{dt} \sigma_n(x - t\mathbf{1}_n).$$

Note that $\eta(x) \in \mathbb{R}^{n-1}$ whenever $x \in \mathbb{R}^n$. By the definition of $\eta(x)$, we obtain the factorization

$$\begin{aligned} \frac{d}{dt} \sigma_n(x - t\mathbf{1}_n) &= (\eta_1(x) - t)(\eta_2(x) - t) \cdots (\eta_{n-1}(x) - t) \\ &= \sigma_{n-1}(\eta(x) - t\mathbf{1}_{n-1}). \end{aligned} \tag{11}$$

Note that on the left, the elementary symmetric polynomial is in n variables, whereas on the right the elementary symmetric polynomial is in $n-1$ variables. This immediately yields a recursive description of $\mathbb{R}_+^{n,(k)}$.

Lemma 1. *If $1 \leq k \leq n-1$ then $\mathbb{R}_+^{n,(k)} = \left\{ x \in \mathbb{R}^n : \eta(x) \in \mathbb{R}_+^{n-1,(k-1)} \right\}$.*

Proof.

$$\begin{aligned} x \in \mathbb{R}_+^{n,(k)} &\iff \text{all roots of } t \mapsto \frac{d^k}{dt^k} \sigma_n(x - t\mathbf{1}_n) \text{ are non-negative} \\ &\stackrel{*}{\iff} \text{all roots of } t \mapsto \frac{d^{k-1}}{dt^{k-1}} \sigma_{n-1}(\eta(x) - t\mathbf{1}_{n-1}) \text{ are non-negative} \\ &\iff \eta(x) \in \mathbb{R}_+^{n-1,(k-1)}. \end{aligned}$$

where the equivalence marked with an asterisk follows from (11). □

Next we show that $\eta(x)$ can be written as the eigenvalues of a symmetric matrix with entries that are linear in x . Our argument is essentially a restatement of Sanyal's spectrahedral representation of $\mathbb{R}_+^{n,(1)}$. Let V_n denote an $n \times (n-1)$ matrix whose columns are an orthonormal basis for the subspace $\{x \in \mathbb{R}^n : \sum_{i=1}^n x_i = 0\}$.

Proposition 6. *If $x \in \mathbb{R}^n$ then*

$$\frac{d}{dt} \sigma_n(x - t\mathbf{1}_n) = \sigma_{n-1}(x - t\mathbf{1}_n) = n \det(V_n^T \text{diag}(x - t\mathbf{1}_n) V_n).$$

Hence $\eta(x) = \lambda(V_n^T \text{diag}(x) V_n)$ and

$$\mathbb{R}_+^{n,(k)} = \left\{ x \in \mathbb{R}^n : V_n^T \text{diag}(x) V_n \in \lambda^{-1}[\mathbb{R}_+^{n-1,(k-1)}] \right\}. \tag{12}$$

Proof. The first assertion is a slight variation on Theorem 1.1 of [17] and follows from the Cauchy-Binet formula. To see that $\eta(x) = \lambda(V_n^T \text{diag}(x)V_n)$ simply note that $V_n^T \text{diag}(x - t\mathbf{1}_n)V_n = V_n^T \text{diag}(x)V_n - tI$ because $V_n^T \text{diag}(\mathbf{1}_n)V_n = V_n^T V_n = I$. The recursive representation of $\mathbb{R}_+^{n,(k)}$ then follows directly from Lemma 1. \square

Expanding the recursive description in (12) gives our root-based representation of $\mathbb{R}_+^{n,(k)}$. The representation as stated in Theorem 2 can be turned into an explicit semidefinite representation by appealing to Proposition 2.

Theorem 2. *Suppose $1 \leq k \leq n-1$. Then $x \in \mathbb{R}_+^{n,(k)}$ if and only if there exist $y_1 \in \mathbb{R}^{n-1}$, $y_2 \in \mathbb{R}^{n-2}$, \dots , $y_k \in \mathbb{R}^{n-k}$ such that*

$$\begin{aligned} \lambda(V_n^T \text{diag}(x)V_n) &\leq y_1 \\ \lambda(V_{n-1}^T \text{diag}(y_1)V_{n-1}) &\leq y_2 \\ &\vdots \\ \lambda(V_{n-k+1}^T \text{diag}(y_{k-1})V_{n-k+1}) &\leq y_k \\ y_k &\geq 0. \end{aligned}$$

This representation is symmetric in the x_i and polynomial-sized. Indeed each eigenvalue majorization condition requires $O(n)$ additional matrix variables each of size $O(n)$ (by Proposition 2) and there are k such majorization conditions for a total size of $O(n^2k)$.

3.2 Coefficient-based representation

A drawback of the root-based representation introduced in Section 3.1 is that it gives very complicated descriptions of $\mathbb{R}_+^{n,(n-k)}$ when k is small. Motivated by this, we now present our coefficient-based representation which gives simpler descriptions of $\mathbb{R}_+^{n,(n-k)}$ when k is small.

In this section we initially work with the open cones $\mathbb{R}_{++}^{n,(n-k)}$ and consider the closure later, because the open cones admit the representation (cf. (4))

$$\mathbb{R}_{++}^{n,(n-k)} = \{x \in \mathbb{R}^n : (\sigma_1/\sigma_0)(x) > 0, (\sigma_2/\sigma_1)(x) > 0, \dots, (\sigma_k/\sigma_{k-1})(x) > 0\}.$$

This description is particularly nice because the functions σ_{k+1}/σ_k are concave on $\mathbb{R}_{++}^{n,(n-k)}$ [16, Theorem 7], and interact well with the operation of taking the polar derivative of a polynomial, which we describe next.

Polar derivatives: Given a univariate polynomial $p(t) = \sum_{k=0}^n a_{n-k}t^k$, of degree n with real roots, let $p^{\text{hom}}(s, t) = \sum_{k=0}^n a_{n-k}s^{n-k}t^k$ denote the corresponding bivariate homogeneous polynomial. Let

$$\mathcal{D}^{\text{polar}} p(t) := \frac{\partial}{\partial s} p^{\text{hom}}(1, t) = \sum_{k=0}^{n-1} (n-k)a_{n-k}t^k$$

denote the *polar derivative* of p . As long as $a_1 \neq 0$, $\mathcal{D}^{\text{polar}}$ is of degree $n-1$. We can also think of the polar derivative as the composition of reversing the coefficients of the polynomial (inverting its roots), differentiating the result, and then reversing the coefficients again (again inverting the roots). As such, if p has real roots, then so does $\mathcal{D}^{\text{polar}} p$.

In the case where $p(t) = \sigma_n(x - t\mathbf{1}_n)$,

$$\mathcal{D}^{\text{polar}} p(t) = \sum_{k=0}^{n-1} (-1)^k (n-k) \sigma_{n-k}(x) t^k. \quad (13)$$

If $\sigma_1(x) \neq 0$ then define

$$\eta^p(x) := \text{roots of the polar derivative of } t \mapsto \sigma_n(x - t\mathbf{1}_n)$$

so that $\mathcal{D}^{\text{polar}} p$ factors as

$$\mathcal{D}^{\text{polar}} p(t) = (-1)^{n-1} \sigma_1(x) (t - \eta_1^p(x)) \cdots (t - \eta_{n-1}^p(x)). \quad (14)$$

If, in addition, $x_i \neq 0$ for all i then by the reverse-differentiate-reverse interpretation of $\mathcal{D}^{\text{polar}}$:

$$\eta^p(x) = \eta(x^{-1})^{-1}$$

where $x^{-1} \in \mathbb{R}^n$ is the element-wise inverse of x .

We now show how quotients of elementary symmetric polynomials and roots of the polar derivative interact. In Lemma 2 below, note that the elementary symmetric polynomials on the left are in n variables and those on the right are in $n-1$ variables.

Lemma 2. *If $1 \leq k \leq n-1$ and $x \in \mathbb{R}^n$ is such that $\sigma_1(x) \neq 0$ then*

$$(\sigma_{k+1}/\sigma_k)(x) = \frac{k}{k+1} \cdot (\sigma_k/\sigma_{k-1})(\eta^p(x)).$$

Proof. By comparing (13) and (14) we see that for $k = 1, 2, \dots, n-1$

$$\begin{aligned} \sigma_k(\eta^p(x)) &= (k+1) \sigma_{k+1}(x) / \sigma_1(x) \quad \text{and} \\ \sigma_{k-1}(\eta^p(x)) &= k \sigma_k(x) / \sigma_1(x). \end{aligned}$$

Computing the appropriate ratio and rearranging yields the identity. \square

Lemma 2 provides us with another recursive description of the derivative relaxations of the orthant.

Lemma 3. *If $1 \leq k \leq n-1$*

$$\mathbb{R}_{++}^{n, (n-k)} = \{x \in \mathbb{R}^n : \sigma_1(x) > 0, \eta^p(x) \in \mathbb{R}_{++}^{n-1, (n-k)}\}.$$

Proof.

$$\begin{aligned} x \in \mathbb{R}_{++}^{n, (n-k)} &\iff (\sigma_1/\sigma_0)(x) > 0, (\sigma_2/\sigma_1)(x) > 0, \dots, (\sigma_k/\sigma_{k-1})(x) > 0 \\ &\iff^* (\sigma_1/\sigma_0)(x) > 0 \text{ and} \\ &\quad (\sigma_1/\sigma_0)(\eta^p(x)) > 0, \dots, (\sigma_{k-1}/\sigma_{k-2})(\eta^p(x)) > 0 \\ &\iff \sigma_1(x) > 0 \text{ and } \eta^p(x) \in \mathbb{R}_+^{n-1, (k-1)}. \end{aligned}$$

where the equivalence marked with an asterisk follows from Lemma 2. \square

We now show that if $\sigma_1(x) > 0$ then $\eta^p(x)$ can be expressed in terms of the eigenvalues of a Schur complement of a symmetric matrix with entries that are linear in x , allowing us to write $\mathbb{R}_+^{n,(n-k)}$ in terms of $\lambda_{sc}^{-1}[\mathbb{R}_+^{n-1,(n-k)}]$. Here, as earlier, Q_n is the orthogonal matrix $Q_n = [\mathbf{1}_n/\sqrt{n} \quad V_n]$.

Proposition 7. *If $1 \leq k \leq n-1$*

$$\mathbb{R}_+^{n,(n-k)} = \left\{ x \in \mathbb{R}^n : Q_n^T \text{diag}(x) Q_n \in \lambda_{sc}^{-1}[\mathbb{R}_+^{n-1,(n-k)}] \right\}$$

Proof. By the definition of $\lambda_{sc}^{-1}[C]$, the statement is equivalent to establishing that $\mathbb{R}_+^{n,(n-k)}$ equals

$$\text{cl} \left\{ x \in \mathbb{R}^n : \sigma_1(x) > 0, \quad \lambda \left(V_n^T \text{diag}(x) V_n - \frac{V_n^T x x^T V_n}{\sigma_1(x)} \right) \in \mathbb{R}_+^{n-1,(n-k)} \right\}.$$

Assume, in addition, that $\sigma_n(x) \neq 0$ so that $\eta^p(x) = \eta(x^{-1})^{-1}$. By Proposition 6 we have that

$$\eta^p(x) = \lambda \left[(V_n^T \text{diag}(x^{-1}) V_n)^{-1} \right] = \lambda \left(V_n^T \text{diag}(x) V_n - \frac{V_n^T x x^T V_n}{\sigma_1(x)} \right). \quad (15)$$

The second equality holds because $(V_n^T \text{diag}(x^{-1}) V_n)^{-1}$ is the inverse of the lower right submatrix in the partitioned matrix

$$\begin{aligned} \begin{bmatrix} \sigma_1(x)/n & x^T V_n/\sqrt{n} \\ V_n^T x/\sqrt{n} & V_n^T \text{diag}(x) V_n \end{bmatrix}^{-1} &= Q_n^T \text{diag}(x^{-1}) Q_n \\ &= \begin{bmatrix} \sigma_1(x^{-1})/n & (x^{-1})^T V_n/\sqrt{n} \\ V_n^T (x^{-1})/\sqrt{n} & V_n^T \text{diag}(x^{-1}) V_n \end{bmatrix} \end{aligned}$$

and so by a property of the Schur complement [19, Theorem 1.2] can be written (after canceling a common factor of n) as

$$(V_n^T \text{diag}(x)^{-1} V_n)^{-1} = V_n^T \text{diag}(x) V_n - V_n^T x x^T V_n / \sigma_1(x).$$

The identity (15) extends to all x such that $\sigma_1(x) > 0$ (removing the assumption $\sigma_n(x) \neq 0$) since $\eta^p(x)$ and the eigenvalues of the Schur complement are continuous for all such x . The result then follows by taking the closure in the characterization of $\mathbb{R}_{++}^{n,(n-k)}$ in Lemma 3. \square

By recursively applying Propositions 3 and 7 we could explicitly write down our coefficient-based representation of $\mathbb{R}_+^{n,(n-k)}$. Its size would be $O(n^2(n-k))$. Rather than do this we note that Example 1 in Section 1.6 gives a concrete illustration of this process.

3.3 Dual cones

If a cone is semidefinitely representable so is its dual cone and there are explicit procedures to take a semidefinite representation for one and produce a semidefinite representation for the other [13, Section 4.1.1]. In this section we note that by using the results of Section 2.4 we can give representations of the dual cones of $\mathbb{R}_+^{n,(k)}$ that inherit the recursive structure of our original root-based and coefficient-based representations.

Recall from Proposition 6 that $\mathbb{R}_+^{n,(k)} = \{x \in \mathbb{R}^n : \mathcal{A}_n(x) \in \lambda^{-1}[\mathbb{R}_+^{n-1,(k-1)}]\}$ where $\mathcal{A}_n(x) = V_n^T \text{diag}(x) V_n$ is an injective linear map from \mathbb{R}^n to \mathcal{S}^{n-1} . In this case it follows from Lewis' result that

$$\begin{aligned} (\mathbb{R}_+^{n,(k)})^* &= \{\mathcal{A}_n^*(M) : M \in \lambda^{-1}[(\mathbb{R}_+^{n-1,(k-1)})^*]\} \\ &= \{\text{diag}(V_n M V_n^T) : M \in \lambda^{-1}[(\mathbb{R}_+^{n-1,(k-1)})^*]\}. \end{aligned}$$

Because $(\mathbb{R}_+^{n-k,(0)})^* = \mathbb{R}_+^{n-k}$ we can apply this result recursively to obtain an analogue of Theorem 2 for the cones $(\mathbb{R}_+^{n,(k)})^*$.

Theorem 3. *Suppose $1 \leq k \leq n-1$. Then $x \in (\mathbb{R}_+^{n,(k)})^*$ if and only if there exist $M_1 \in \mathcal{S}^{n-1}, M_2 \in \mathcal{S}^{n-2}, \dots, M_k \in \mathcal{S}^{n-k}$ such that*

$$\begin{aligned} x &= \text{diag}(V_n M_1 V_n^T) \\ \lambda(M_1) &\preceq \text{diag}(V_{n-1} M_2 V_{n-1}^T) \\ &\vdots \\ \lambda(M_{k-1}) &\preceq \text{diag}(V_{n-k+1} M_k V_{n-k+1}^T) \\ M_k &\succeq 0 \end{aligned}$$

To similarly dualize the coefficient based representation, recall from Proposition 7 that $\mathbb{R}_+^{n,(n-k)} = \{x \in \mathbb{R}^n : \mathcal{B}_n(x) \in \lambda_{sc}^{-1}[\mathbb{R}_+^{n-1,(n-k)}]\}$ where $\mathcal{B}_n(x) = Q_n^T \text{diag}(x) Q_n$ is an injective linear map from \mathbb{R}^n to \mathcal{S}^n . Hence

$$\begin{aligned} (\mathbb{R}_+^{n,(k)})^* &= \{\mathcal{B}_n^*(M) : M \in (\lambda_{sc}^{-1}[\mathbb{R}_+^{n-1,(n-k)}])^*\} \\ &= \{\text{diag}(Q_n M Q_n^T) : M \in (\lambda_{sc}^{-1}[\mathbb{R}_+^{n-1,(n-k)}])^*\} \end{aligned}$$

which can be expanded into an explicit semidefinite representation using Proposition 5.

4 Derivative relaxations of spectrahedral cones

We now show how to use our polynomial-sized semidefinite representations of the cones $\mathbb{R}_+^{m,(k)}$ to construct polynomial-sized semidefinite representations for derivative relaxations of arbitrary spectrahedral cones. Recall from Proposition 1 that if p has degree m and is hyperbolic with respect to e we can express the derivative cone $\Lambda_+^{(k)}(p, e)$ in terms of the cone $\mathbb{R}_+^{m,(k)}$ via the characteristic roots $\lambda_e(x)$. If we can represent $\lambda_e(x)$ in terms of the eigenvalues of a symmetric matrix with entries linear in x , then we can use our semidefinite representation of $\mathbb{R}_+^{m,(k)}$ to give a semidefinite representation of $\Lambda_+^{(k)}(p, e)$. This is precisely the situation we are in when $\Lambda_+(p, e)$ is spectrahedral.

Proposition 8. *Let $A_1, A_2, \dots, A_n \in \mathcal{S}^m$ and $e \in \mathbb{R}^n$ be such that $\sum_{i=1}^n A_i e_i =: B \succ 0$. Let $p(x) = \det(\sum_{i=1}^n A_i x_i)$. Then for $1 \leq k \leq m-1$*

$$\Lambda_+^{(k)}(p, e) = \left\{ x \in \mathbb{R}^n : \exists y \in \mathbb{R}_+^{m,(k)} \text{ s.t. } \lambda(B^{-1/2}(\sum_{i=1}^n A_i x_i)B^{-1/2}) \preceq y \right\}.$$

Proof. Note that $\lambda_e(x) = \lambda(B^{-1/2}(\sum_{i=1}^n A_i x_i)B^{-1/2})$ and so is given by the eigenvalues of a symmetric matrix linearly parameterized by x . The result follows by combining Proposition 1 and (6). \square

By also appealing to Proposition 4 a similar argument could be used to give a semidefinite representation for the dual cone $(\Lambda_+^{(k)}(p, e))^*$ in terms of a semidefinite representation for $(\mathbb{R}_+^{m,(k)})^*$. We omit the details.

We conclude this section with an example of these constructions.

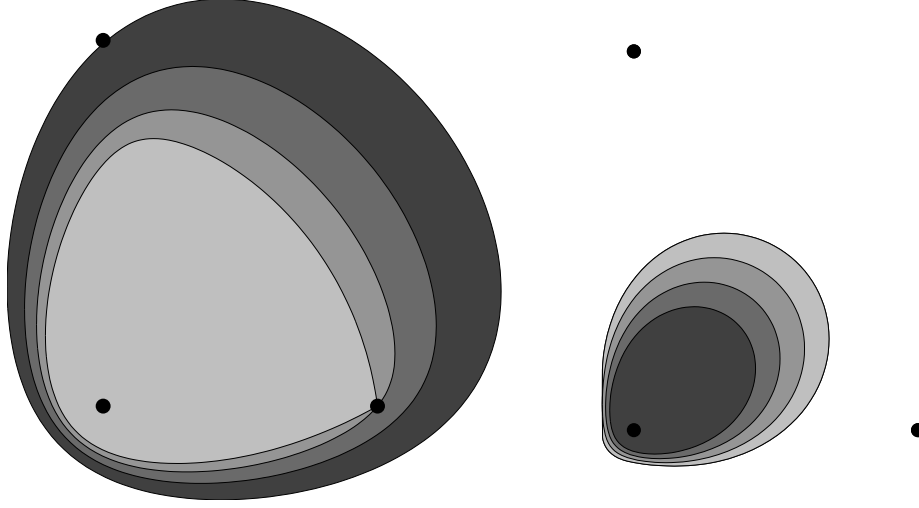


Figure 1: On the left, the inner region is the 3-ellipse consisting of points with sum-of-distances to $(0,0)$, $(0,4)$, and $(3,0)$ equal to 8, i.e. the $z = 1$ slice of the spectrahedral cone defined by (16). The outer three regions are the $z = 1$ slices of the first three derivative relaxations of this spectrahedral cone in the direction $(0,0,1)$. On the right are the $z = 1$ slices of the dual cones of the cones shown on the left, with dual pairs having the same shading.

Example 2 (Derivative relaxations of a 3-ellipse). Given foci $(0,0)$, $(0,4)$ and $(3,0)$ in the plane, the 3-ellipse consisting of points such that the sum of distances to the foci equals 8 is shown in Figure 1. This is one connected component of the real algebraic curve of degree 8 given by $\{(x,y) \in \mathbb{R}^2 : \det E(x,y,1) = 0\}$ where E is defined in (16) (see Nie et al. [14]). The region enclosed by this 3-ellipse is the $z = 1$ slice of the spectrahedral cone defined by $E(x,y,z) \succeq 0$ where

$$E(x,y,z) = \begin{bmatrix} 5z+3x & y & y-4z & 0 & y & 0 & 0 & 0 \\ y & 5z+x & 0 & y-4z & 0 & y & 0 & 0 \\ y-4z & 0 & 5z+x & y & 0 & 0 & y & 0 \\ 0 & y-4z & y & 5z-x & 0 & 0 & 0 & y \\ y & 0 & 0 & 0 & 11z+x & y & y-4z & 0 \\ 0 & y & 0 & 0 & y & 11z-x & 0 & y-4z \\ 0 & 0 & y & 0 & y-4z & 0 & 11z-x & y \\ 0 & 0 & 0 & y & 0 & y-4z & y & 11z-3x \end{bmatrix}. \quad (16)$$

Note that $E(0,0,1) \succ 0$ and so $e = (0,0,1)$ is a direction of hyperbolicity for $p(x,y,z) = \det E(x,y,z)$. The left of Figure 1 shows the $z = 1$ slice of the cone $\Lambda_+(p, e)$ and its first three derivative relaxations $\Lambda_+^{(1)}(p, e)$, $\Lambda_+^{(2)}(p, e)$, and $\Lambda_+^{(3)}(p, e)$. The right of Figure 1 shows the $z = 1$ slice of the cones $(\Lambda_+(p, e))^*$, $(\Lambda_+^{(1)}(p, e))^*$, $(\Lambda_+^{(2)}(p, e))^*$, and $(\Lambda_+^{(3)}(p, e))^*$. All of these convex bodies were plotted by computing 200 points on their respective boundaries by optimizing 200 different linear functionals over them. We performed the optimization by modeling our semidefinite representations of these cones in YALMIP [12] which numerically solved the corresponding semidefinite program using SDPT3 [18].

5 Concluding remarks

We conclude with some comments about (the possibility of) simplifying our representations and some open questions.

5.1 Simplifications

If we can simplify a representation of $\mathbb{R}_+^{n,(k)}$ for some $k = i$, that allows us to simplify the root-based representations for $k \geq i$ and the coefficient-based representations for $k \leq i$. For example $\mathbb{R}_+^{n,(n-2)}$ can be succinctly expressed in terms of the second-order cone $\mathcal{Q}_+^{n+1} = \{x \in \mathbb{R}^{n+1} : (\sum_{i=1}^n x_i^2)^{1/2} \leq x_{n+1}\}$ as

$$\mathbb{R}_+^{n,(n-2)} = \{x \in \mathbb{R}^n : (x, \sigma_1(x)) \in \mathcal{Q}_+^{n+1}\}.$$

Then we can represent $\lambda^{-1}[\mathbb{R}_+^{n,(n-2)}]$ in terms of the second order cone as

$$\lambda^{-1}[\mathbb{R}_+^{n,(n-2)}] = \{M \in \mathcal{S}^n : (M, \text{tr}(M)) \in \mathcal{Q}_+^{(n^2+1)}\}$$

because $\text{tr}(M) = \sum_{i=1}^n \lambda_i(M)$ and $\sum_{i,j=1}^n M_{ij}^2 = \sum_{i=1}^n \lambda_i(M)^2$. Consequently Propositions 3 and 7 can be used to give a concise representation of $\mathbb{R}_+^{n,(n-3)}$ in terms of the second order cone as

$$x \in \mathbb{R}_+^{n,(n-3)} \iff \exists M \in \mathcal{S}^{n-1} \text{ such that} \\ Q_n^T \text{diag}(x) Q_n \succeq \begin{bmatrix} 0 & 0 \\ 0 & M \end{bmatrix} \text{ and } (M, \text{tr}(M)) \in \mathcal{Q}_+^{(n-1)^2+1}.$$

This should be used as a base case instead of $\mathbb{R}_+^{n,(n-1)}$ in the coefficient-based representations. This would, for example, considerably simplify Example 1. Similarly $\mathbb{R}_+^{n,(1)}$ can be expressed as

$$\mathbb{R}_+^{n,(1)} = \{x \in \mathbb{R}^n : V_n^T \text{diag}(x) V_n \succeq 0\}$$

which should be used as the base case instead of \mathbb{R}_+^n in the root-based representations.

5.2 Lower bounds on the size of representations

The explicit constructions given in this paper establish upper bounds on the minimum size of semidefinite representations of $\mathbb{R}_+^{n,(k)}$. To assess how good our representations are, it is interesting to establish corresponding *lower* bounds on the size of semidefinite representations of $\mathbb{R}_+^{n,(k)}$. In the case of $\mathbb{R}_+^{n,(n-1)}$, a halfspace, the obvious semidefinite representation of size one is clearly of minimum size. Less trivial is the case of $\mathbb{R}_+^{n,(0)}$, the non-negative orthant. It has been shown by Gouveia et al. [5, Section 5] that \mathbb{R}_+^n does not admit a semidefinite representation of size smaller than n . Hence the obvious representation of \mathbb{R}_+^n as the restriction of \mathcal{S}_+^n to the diagonal is of minimum size. We are not aware of any non-trivial lower bounds on the size of semidefinite representations of the cones $\mathbb{R}_+^{n,(k)}$ for $1 \leq k \leq n-2$.

The semidefinite representations of $\mathbb{R}_+^{n,(k)}$ given in this paper are *equivariant* in that they appropriately preserve the symmetries of $\mathbb{R}_+^{n,(k)}$. (For a precise definition see [5, Definition 2.10].) It is known that symmetry matters when representing convex sets as projections of other convex sets [9]. For example if p is a power of a prime, equivariant representations of regular p -gons in \mathbb{R}^2 are necessarily much larger than their minimum-sized non-equivariant counterparts [5, Proposition

3.5]. Given that the cones $\mathbb{R}_+^{n,(k)}$ are highly symmetric, it would also be interesting to establish lower bounds on the size of equivariant semidefinite representations of the derivative relaxations of the non-negative orthant.

References

- [1] V. Anh Truong and L. Tunçel. Geometry of homogeneous convex cones, duality mapping, and optimal self-concordant barriers. *Math. Program.*, 100(2):295–316, 2004.
- [2] P. Brändén. Hyperbolicity cones of elementary symmetric polynomials are spectrahedral. arXiv:1204.2997, 2012.
- [3] Y.-B. Choe, J. G. Oxley, A. D. Sokal, and D. G. Wagner. Homogeneous multivariate polynomials with the half-plane property. *Adv. Appl. Math.*, 32(1–2):88–187, 2004.
- [4] J. Gouveia and T. Netzer. Positive polynomials and projections of spectrahedra. *SIAM J. Optim.*, 21(3):960, 2011.
- [5] J. Gouveia, P.A. Parrilo, and R. Thomas. Lifts of convex sets and cone factorizations. arXiv:1111.3164, 2011.
- [6] L. Gårding. An inequality for hyperbolic polynomials. *J. Math. Mech.*, 8(6):957–965, 1959.
- [7] O. Güler. Hyperbolic polynomials and interior point methods for convex programming. *Math. Oper. Res.*, 22(2):350–377, 1997.
- [8] J.W. Helton and V. Vinnikov. Linear matrix inequality representation of sets. *Comm. Pure Appl. Math.*, 60(5):654–674, 2007.
- [9] V. Kaibel, K. Pashkovich, and D. Theis. Symmetry matters for the sizes of extended formulations. In *Integer Programming and Combinatorial Optimization*, volume 6080 of *Lecture Notes in Comput. Sci.*, pages 135–148. Springer Berlin, 2010.
- [10] A. S. Lewis. Convex analysis on the hermitian matrices. *SIAM J. Optim.*, 6(1):164–177, 1996.
- [11] A.S. Lewis, P.A. Parrilo, and M.V. Ramana. The Lax conjecture is true. *Proc. Amer. Math. Soc.*, 133(9):2495–2500, 2005.
- [12] J. Löfberg. YALMIP: A toolbox for modeling and optimization in MATLAB. In *Proceedings of the CACSD Conference*, Taipei, Taiwan, 2004.
- [13] A. Nemirovski. Advances in convex optimization: conic programming. In *Proceedings of the International Congress of Mathematicians: Madrid, August 22–30, 2006: invited lectures*, pages 413–444, 2006.
- [14] J. Nie, P.A. Parrilo, and B. Sturmfels. Semidefinite representation of the k -ellipse. In *Algorithms in Algebraic Geometry*, volume 146 of *The IMA Volumes in Mathematics and its Applications*, pages 117–132. Springer New York, 2008.
- [15] J. Renegar. Hyperbolic programs, and their derivative relaxations. *Found. Comput. Math.*, 6:59–79, 2006.
- [16] J. Renegar. Central swaths (a generalization of the central path). arXiv:1005.5495, 2010.

- [17] R. Sanyal. On the derivative cones of polyhedral cones. *Adv. Geom.*, 2012. to appear.
- [18] K.C. Toh, M.J. Todd, and R.H. Tütüncü. SDPT3—a MATLAB software package for semidefinite programming, version 1.3. *Optim. Methods Softw.*, 11(1-4):545–581, 1999.
- [19] F. Zhang, editor. *The Schur complement and its applications*. Springer, 2005.
- [20] Y. Zinchenko. On hyperbolicity cones associated with elementary symmetric polynomials. *Optim. Lett.*, 2(3):389–402, 2008.